Techniques for Text, Image, Audio and Video Indexing and Retrieval

Ms. Nayana S. Zope[1], Mr. Amit Gadekar[2]

[1] Pune University, Student SITRC Nashik,

[2] Assistant Professor of Computer Engineering, Pune University,

Abstract

The information retrieval is an interdisciplinary field. The Retrieval of multimedia has become a requirement for many contemporary systems. These systems provide to searching, querying, and navigation and sometimes composition capabilities involving various forms of media. In this survey paper, we review the techniques for text, image, audio and video retrieval. We focus on indexing and retrieval techniques for text, image, audio and video. Here also discuss features visual features for video retrieval such as colour, texture, shape. The indexing techniques are discussed for these features. We also compare most popular techniques used for indexing and retrieval.

Keywords: Information Retrieval, Extraction, Indexing, Retrieval, Multimedia.

1. INTRODUCTION

Information Retrieval (IR) our basic task is to find the subset of a collection of elements that is relevant to a query. An information retrieval process begins when a user enters a query into the system. Although the search engines are depends on rank websites based on some combination of their popularity. They are not all the same but describes to crawler-based search engines. Each type of "search engines" gathers and ranks listings in radically different ways. Crawler-based search engines such as Google, Yahoo etc., compile their listing automatically. The people see their listings search through the "Crawl" or "Spider" the web. These listings are what make up the search engine's index or catalogue. It can think of the index as a massive electronic filing cabinet containing a copy of every web page the spider i.e. content index finds. Because spiders scour the web on a regular basis, any changes made to web site may affect search engine ranking. It is also important to remember that it make take a while for a spidered page added to index. It is not available to those searching with the search engine until that happens.

Directories such as open directory depend on human editors to compile their listing. Web masters submit an address, title and a brief description of their site and submit the review. The hybrid search engines will typically favour one type of listing over the however.

Rest of paper is organised as Section 2 reviews various algorithms and approaches used for text informational retrieval, Section 3 we explain the various indexing and retrieval techniques fir image, Section 4 discusses audio indexing retrieval techniques, Section 5 deals with techniques for video indexing and retrieval and finally in Section 6 we conclude.

2. TEXT RETRIEVAL AND INDEXING

The basic document retrieval process is as shown in figure 1. As shown on the right side of the figure, the documents are processed off-line to obtain document representations. These representations are stored together with documents themselves. During retrieval, the user issues a query that is processed (online) to obtain its representation. Then the query representation is compared with the document representation. Document deemed relevant by the system are retrieved and presented by user, who evaluates the returned documents and decides which ones are actually relevant to the information need. A good IR system should then allow the user to provide relevance feedback to the system.
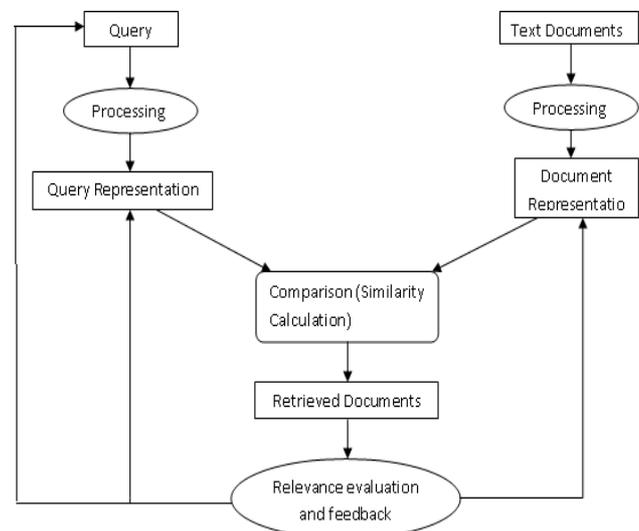

Figure 1 Basic Text Retrieval model

The system uses this information to modify query, query representation, and/or document representation. Retrieval is done based o n the modified query and document representation. If necessary, the retrieval feedback process is iterated a few times. Note that not all IR systems support the user relevance feedback process. Following Table 1 briefly review various retrieval approaches with its advantages and limitation.

Table 1 Comparison of Text Retrieval and Indexing Approaches

.Algorithm	.Approaches	.Advantages	.Limitation
Basic Boolean	Documents are indexed by set of keywords. Queries are also represented by a set of keywords joined by logical operators that supply relationships between the query terms.	Simple to implement. Ideal for small text data patterns and databases.	It is difficult to formulate Boolean queries and the retrieval results are very sensitive to query formulation, Query term weights are normally not used as queries are often very short.
Vector Space Model Indexing and Retrieval		When subsequent queries similar to the queries used to modify the documents are issued, performance is improved.	Treats terms as unrelated and it works only well short documents and queries.
Probabilistic Based Indexing and Retrieval	It is based on the following four parameters: P(rel):The probability of relevance of a document P(nonrel): The probability non relevance of a document a1 : The cost associated with the retrieval of a non relevant document a2: The cost associated with the non retrieval of a relevant document	Since the retrieval of a non relevant document carries a loss of a1 P(nonrel) and the rejection of a relevant document carries a loss of a2P(rel), the total loss caused by a given retrieval process will be minimised if a document is retrieved whenever $a_2P(rel) >= a_1P(nonrel)$	Retrieval effectiveness and efficiency are low because not all relevant items may be retrieved and whole document space has to be searched.
Cluster Based Indexing and Retrieval	Document clustering- grouping similar documents into cluster are used. During	Effective Retrieval	High computation, memory and time requirement

document retrieval, the query vector is compared with the controls of cluster. After the cluster with highest similarity to the query vector is identified, either all documents in the cluster are retrieved or most similar documents are retrieved.

3. IMAGE INDEXING AND RETRIEVAL

In Image indexing and retrieval section describes five main approaches, which are based on structured attributes, object recognition, text and low level image features. The first approaches, attribute based, uses traditional database management systems for image indexing and retrieval. The second approach colour and third based on shape are not mature yet as it relies in automatic object recognition. Text-Based image retrieval uses traditional IR for image indexing and retrieval. The low-level content based approach to image indexing and retrieval queries the extraction of low level image features. The common features used are colour, object, shape and texture. Table 2 looks like into a number of image indexing and retrieval techniques based on compressed image data.

Table 2 Comparison of Image Indexing and Retrieval Approaches

.Algorithm	.Approaches	.Advantages	.Limitation
Text-Based Indexing and Retrieval Techniques	Manual, Automated, Annotation, Attributes based indexing and retrieval	Easy to implement, low memory and computational requirement	Most text-based retrieval systems require manual annotation of images.
Image Retrieval Based on Shape	Histograms, Spatial relation among pixels, colour distribution etc.	Histogram is easy to compute and effective in characterising both global and local distribution of colour in an image.	Retrieval performance may be poor.
Image Retrieval Based on	Major and Minor axis, rectangle,	Object can be recognised and retrieved.	

International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)
Web Site: www.ijettcs.org Email: editor@ijettcs.org
Volume 4, Issue 5(2), September - October 2015   **ISSN 2278-6856**

Shape	Eccentricity and moment invariants, histogram of edges detected, elastic temple matching, region based shape representation, Rotational Normalisation.		
Sequential Multiple Attribute Tree	The SMAT is a multi-tier tree structure, where each layer corresponds to an indexing attribute.	Multiple features are used for indexing.	
Texture Based Methods	Coarseness, Contract, Directionality, Line Likeness, Regularity, Roughness	Ideal for natural Images and satellite Images.	Poor performance on images other than natural images.

4. AUDIO INDEXING AND RETRIEVAL

Improvements in speech recognition technology and computing power have enabled the development of usable indexes for vast spoken audio repositories. A standard technique is to use speech recognition to transcribe the audio and then to build an index using this transcription. However, this approach suffers from that fact that a speech recogniser has a limited vocabulary so the system cannot retrieve out of vocabulary (OOV) queries. A popular technique to confront this problem is to use phoneme recognizition system is used to transcribe the spoken audio. Word queries are then converted to phoneme sequences and searched for in the transcription.

Table 3 Comparison of Audio Indexing and Retrieval Approaches

.Algorithm	.Approaches	.Advantages	.Limitation
Techniques Based on Artificial Neural Network	An ANN consists of many neurons interconnected by links with weights. Speech recognition with ANNs consists of training and recognition stages.	Suitable speech recognition	Extensive training is required for effective retrieval
Virtual-Node	It exploits the	Fast and	High

(Vn) Algorithm	multi-resolution capabilities of the hierarchical tree structure	effective retrieval	memory requirement
Feature Vector based technique	Pitch, ZCR etc. features are extracted and indexed	Good retrieval result	Require more time to compare query with database
Techniques Based on Artificial Neural Network	An ANN consists of many neurons interconnected by links with weights. Speech recognition with ANNs consists of training and recognition stages.	Suitable speech recognition	Extensive training is required for effective retrieval
Virtual-Node (Vn) Algorithm	It exploits the multi-resolution capabilities of the hierarchical tree structure	Fast and effective retrieval	High memory requirement

5. VIDEO INDEXING AND RETRIEVAL

With the development of multimedia data types and available bandwidth there is huge demand of video retrieval systems. Content based Video Indexing and Retrieval (CBVIR), in this problem searching for digital videos in large databases. The "Content-Based" means the search will analyse the actual content of the video. Video is created by taking set of shots and composing them together specified composition operators. Due to limitation of space in this paper we discuss texture, colour and shape used for indexing.

Table 4 Comparison of Video Indexing and Retrieval Approaches

.Algorithm	.Approaches	.Advantages	.Limitation
Indexing and retrieval Based on r frames of video shots	Common way of creating a shot index is to use a representative frame to represent a shot.	Single frame represent a shot, reducing complex computation.	Treats video as a collection of images and ignores temporal or motion information contained in the video
Indexing and retrieval Based on Motion Information	Motion content, motion uniformity etc. is	Extension of r frame method and takes care of motion	Retrieval performance depends on window size.

	captured from video.	information	
Indexing and retrieval Based on Metadata	Video metadata is used for indexing	Suitable for conventional DBMSs	
Indexing and retrieval Based on Annotation	Manual or automated annotation	Performs well if transcripts are available	This approach is very challenging because speech and no speech are normally
Indexing and retrieval Based on r frames of video shots	Common way of creating a shot index is to use a representative frame to represent a shot.	Single frame represent a shot, reducing complex computation.	Treats video as a collection of images and ignores temporal or motion information contained in the video

6. CONCLUSION

Despite the considerable progress of academic research in multimedia retrieval, there has been relatively little impact on content based video retrieval research on commercial applications with some niche exceptions such as video segmentation. One promising approach is to use meat learning to automatically select or combine appropriate features. The possibility is to develop an interactive user interface based on visually interpreting the data using a selected measure to assist the selection process. Also the user interactions are indispensable in the determination of features, to develop new theories, methods and tools to facilitate the user's involvement. In the literature, a large number of algorithms have been presented. The different information sources (e.g. colour, texture, motion, shape, geometry, etc.) are used for text, audio video and images. By merging the different sources of information we and enhance the performance of all text, audio, video and image extraction system and retrieval systems.

References

[1] S. M.Metev ankanhalli and V. P. Kankanhalli and Y. Rui, "Application Potentail of Multimedia Information Retrieval", Proc. IEEE, April 2008.

[2] Sinha and Ramesh Jain, "Concept Annotation and research space Decrement of Digital Photos using Optical Content Information", in proceeding of SPIE, Multimedia content access: Algorithms and system, January 2008.

[3] Jiawei Han and Micheline Kamber, "Data Mining: concepts technique", 2nd edition, Morgan Kaufmann Publishers, San Francisco, 2006.

[4] Keiji Yanai, Kobus Barnard, "Finding Visual Concepts by Web Image MInign", in Proc. Of www 2006, May 23-26, 2006, Edinburgh, Scotland.

[5] Jung, Keechul, Kwang in Kim, and Anil K Jain. "Text information extraction in images and video: a survey." Pattern recognition 37, no. 5 (2004): 977-997.

[6] Y. K. Lim, S. H. Choi, and S.W. Lee, "Text Extraction in MPEG Compressed Video for Content-based Indexing", Proc. of International Conference on Pattern Recognition, 2000 pp. 409-412.

AUTHOR

Nayana S. Zope received the B.E. degree in Information Technology in North Maharashtra University and M.E. (Appear) degree in Computer Engineering from Pune University in 2006 and 2015, respectively. During 2011-2015, she stayed in Godavari Foundation as a teaching organization to study Data Warehousing and Mining, Database Management System, Information Retrieval, Programming Laboratory-I, Advanced Development Tools and Laboratory etc.